```
=> d full his

     (FILE 'HOME' ENTERED AT 11:06:59 ON 02 APR 2003)

     FILE 'REGISTRY' ENTERED AT 11:15:56 ON 02 APR 2003
L1             1 SEA ABB=ON  PLU=ON  9000-83-3/RN

     FILE 'HCAPLUS' ENTERED AT 11:16:09 ON 02 APR 2003

     FILE 'REGISTRY' ENTERED AT 11:16:13 ON 02 APR 2003
                SET SMARTSELECT ON
L2              SEL PLU=ON  L1 1- CHEM :      12 TERMS
                SET SMARTSELECT OFF

     FILE 'HCAPLUS' ENTERED AT 11:16:14 ON 02 APR 2003
L3         76747 SEA ABB=ON  PLU=ON  L2
L4           602 SEA ABB=ON  PLU=ON  L3 (L) (MYOSIN HEAVY CHAIN)
L5            75 SEA ABB=ON  PLU=ON  L4 (L) (MAN OR HUMAN)
L6            50 SEA ABB=ON  PLU=ON  L5 AND PD<19981105
L7             7 SEA ABB=ON  PLU=ON  L6 AND (DNA OR CDNA OR NUCLEIC ACID OR
                 NUCLEOTIDE OR POLYNUCLEOTIDE)
```

WEST Search History

DATE: Wednesday, April 02, 2003

Set Name side by side	Query	Hit Count	Set Name result set
	DB=USPT,PGPB; PLUR=YES; OP=ADJ		
L14	L13 and l8	60	L14
L13	L12 and vector and host	141	L13
L12	L11 and (@ad<19981105)	230	L12
L11	L10 and (dna or cdna or nucleic acid or nucleotide or polynucleotide)	500	L11
L10	L9 and (human or man)	521	L10
L9	myosin heavy chain	530	L9
L8	L7 or l6 or l5 or l4 or l3 or l2 or l1	20386	L8
L7	((536/23.2)!.CCLS.)	6498	L7
L6	(((435/320.1)!.CCLS.))	16086	L6
L5	(((435/252.33)!.CCLS.))	2326	L5
L4	(((435/252.3)!.CCLS.))	6644	L4
L3	(((435/196)!.CCLS.))	671	L3
L2	(((435/195)!.CCLS.))	487	L2
L1	((435/183)!.CCLS.)	2623	L1

END OF SEARCH HISTORY

```
OM of: q39157 to: US09830914C.seq:*   out_format : pfs

Date: Wed Apr  2 14:15:01 2003

About: Results were produced by the GenCore software, version 5.1.4_p5_4578,
       Copyright (c) 1993-2003 Compugen Ltd.

Command line parameters:
-MODEL=frame+_p2n.model -DEV=soft -Q=q39157.pep -DB=US09830914C.seq
-SUFFIX=pfs -OUT=compare -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits
-START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY
-NEG_SCORES=0 -LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-FGAPEXT=0.1 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=0.1

Search information:
Query: q39157
Query length: 1101
Database: US09830914C.seq:*
Database sequences: 1
Database length: 2109
Search time (sec): 2

Score list:
Sequence              Strd  Orig   ZScore   EScore   Len ! Documentation  ..
US09830914C.seq:US-09-830-914C-2 +   850.51    0.00      0.0  2109 !  Sequence 2, Applica
US09830914C.seq:US-09-830-914C-2 -   416.20    0.00      0.0  2109 !  Sequence 2, Applica

Sequence name: US09830914C.seq:US-09-830-914C-2

Sequence documentation:
; Sequence 2, Application US/09830914C
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMIC, INC.; TANG, Y. Tom;
;  APPLICANT:  CORLEY, Neil C.; GORGONE, Gina A.;
;  APPLICANT:  GUEGLER, Karl J.; BAUGHN, Mariah R.
;  TITLE OF INVENTION: MYOSIN HEAVY CHAIN HOMOLOG
;  FILE REFERENCE: PF-0621 USN
;  CURRENT APPLICATION NUMBER: US/09/830,914C
;  CURRENT FILING DATE:  2002-11-06
;  PRIOR APPLICATION NUMBER: PCT/US99/26177
;  PRIOR FILING DATE: 1999-11-05
;  PRIOR APPLICATION NUMBER: US 60/172,248
;  PRIOR FILING DATE: 1998-11-05
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PERL Program
; SEQ ID NO 2
;   LENGTH: 2109
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 1929760CB1
US-09-830-914C-2

Alignment of: q39157 x US-09-830-914C-2   ..

Alignment segment 1/1: (+)

                  Quality:  850.51                    Escore:       0
           Matching length:     577                Total length:     885
 Matching Percent Similarity:   75.56   Matching Percent Identity:   44.54
   Total Percent Similarity:   49.27     Total Percent Identity:   29.04
                    Gaps:      11

Alignment:

     210 Cys...GlyValGluTyrGluIleLeuLysThrThrCysIleLeu.GluA 225
           |||   |||::::::  ::::::::  ::::::|||:::||| |||
       4 TGCTCTGGGCTGTCTTCACACTTCATTTGGGTTTCCTGCTTGCTCTGAGC 53

     225 laPhe..GlyAsnAlaLysThrSerArgAsnAlaAsnSerSerArgPheG 241

         :::  ||||||:::
      54 TCTACAGGGGAATGGG..................................... 69

     241 lyLysLeuIleGluIle.....HisPheSer...AlaMetGlyLysIleC 255
                 :::|||:::     |||::::::      :::||||||:::|
      70 ........GTAGAGATGGGAGCCACCTTGGGTGGAGGGTGGGGAAGGTAT 111

     255 ysGlyAlaLysLeuGluThrPheLeuLeuGluLysSerArgValValGln 271
         ||:::|||
     112 GTTCTGCC.............................................. 119

     272 LeuPheAsnGlyGluArgSerTyrHisIlePheTyrGluLeuCysAlaGl 288
                                            ::::::   |||   ::
     120 ..............................CACCACAGGTGTCATGC 136

     288 yAlaSer.ProIleLeuLysGluArgLeuLysLeuLysThrAlaSerGlu 304
         :   ||| |||
     137 TCACTCAGCCT.......................................... 147

     305 TyrThrTyrLeuSerGlnSerAspCysLeuThrIleAlaGlyValAspAs 321
                                                        ||
     148 ...............................................GA 149

     321 pAlaGln......LysPheHisLysLeuLeuGluAla.............. 331
         |||||||      |||:::      |||      |||
     150 TGCCCAGGCTGCCAAGTATAAGGCATTGGCAGGGGCCCAGCCACCCTGGG 199

     332 PheAspIleVal.GlnIleProLysGluHisGlnGluArgAlaPheAlaL 348
         |||   :::|||    ||||||                           |
     200 TTCC..TTGGTCCCCTATTCCCCAT.......................C 223

     348 euLeuAlaAlaValLeu...TrpLeu.................GlyAsnVa 358
         |||||   ||||||      |||:::                 |||
     224 TGCTCCCTGCAGTGGCCCCATGGGTTCTCTGCCATCTTCCCAGGC..... 268

     358 lSerPheArgValThrAspAsnGluAsnHisValGluValValAlaAspG 375

     268 ..................................................... 268

     375 luAlaValAlaAsnAlaAlaMetLeuMetGlyCysAsnThrGluGluLeu 391

     268 ..................................................... 268

     392 MetValValLeuSerThrArgLysLeuGlnAlaGlyThrAspCysIleAl 408

     268 ..................................................... 268

     408 aLysLysLeuThrLeuArgGlnAlaThrAspMetArgAspGlyIleAlaL 425

     268 ..................................................... 268

     425 ysPheIleTyrAlaAsnLeuPheAspTrpLeuValGluGlnIleAsnIle 441

     268 ..................................................... 268

     442 AlaLeuGluValGlyLysSerArgThrGlyArgSerIleSerIleLeuAs 458
                                                  :::|||||
     269 ...................................CTGCTGGA 276

     458 pIleTyrGlyPheGluSerPheLysAsnAsnSerPheGluGlnPheCysI 475
         |:::|||||||||||||||||||   :::||||||:::||||||:::||||
     277 TGTGTATGGATTTGAATCATTTCCTGACAACAGTCTGGAACAGTTGTGCA 326

     475 leAsnTyrAlaAsnGluArgLeuGlnGlnHisPheAsnArgHisLeuPhe 491
         ||||||||||||||||||:::|||||||||||||||||      |||   :::
     327 TCAACTACGCCAATGAGAAGCTGCAGCAGCATTTTGTGGCTCACTACCTA 376

     492 LysLeuGluGlnGluGluTyrGluGluAspGlyIleAspTrpThrLysVa 508
         :::   :::||||||||||||      :::|||:::::::|||:::   ::
     377 AGGGCCCAGCAGGAGGAATACGCAGTTGAGGGCCTGGAGTGGTCATTCAT 426

     508 lGluPheValAspAsnGlnGluCysLeuAspLeuIleGluLysLysProI 525
         :::::::   |||||||||   ||||||||||||||||||   :::||||
```

```
 427 CAACTACCAGGACAACCAGCCCTGTTTGGATCTCATTGAGGGAAGCCCCA 476

 525 leGlyLeuLeuSerLeuLeuAspGluGluSer......AsnPheProLys 539
     ||:::::   ||||||:::::|||||          |||   |||:::
 477 TCAGCATCTGCTCCCTCATAAATGAGGAATGCCGCCTCAATCGACCCAGC 526

 540 AlaThrAspLeuThrPheAlaAsnLysLeuLysGlnHisLeuLysThr.. 555
     :::              |||::::::|||:::   ::::::::::|||
 527 AGC.............GCACGCCAGCTCCAGACACGCATTGAGACTGC 561

 556 ..........AsnSerCysPhe.........LysGlyGluArgGlyArgA 566
             :::   |||:::         |||   :::|||      :
 562 CCTGGCAGGCAGCCCCTGCCTGGGCCACAATAAGCTCAGCCGGGAGCCCA 611

 566 laPheArgValAsnHisTyrAlaGlyGluValLeuTyrAspThrAsnGly 582
     ::|||   |||   ||||||||||||   |||   |||   |||   |||
 612 GCTTCATTGTGGTGCATTATGCGGGGCCTGTGCGGTACCACACAGCAGGC 661

 583 PheLeuGluLysAsnArgAspProLeuProAlaAspLeuIleAsnLeuLe 599
     ::::::||||||||||:::|||||||:::|||   :::|||   :::||||||
 662 CTGGTGGAGAAGAACAAGGACCCTATCCCACCTGAGCTGACCAGGCTCCT 711

 599 u...SerSerCysAspCysGlnLeuLeuLysLeuPheSerThr..... 612
     |   :::|||   |||      |||:::   ||||||   |||
 712 GCAGCAATCCCAGGACCCCCTGCTCATGGGGCTGTTTCCTACTAACCCCA 761

 613 LysMetArg.................GlyLysSerGlnLysProLeuMet 623
     |||   |||                |||:::|||:::   |||::::::
 762 AAGAGAAGACCCAGGAGGAACCCCCTGGCCAGAGCAGGGCCCCTGTGTTG 811

 624 LeuSerAspSerThrAsnGlnThrValGlyThrLysPheLysGlyGlnLe 640
                       ||||||   :::||||||||||:::::||
 812 ....................ACCGTGGTGTCCAAGTTCAAGGCCTCACT 840

 640 uPheLysLeuMetAsnLysLeuGluAsnThrSerProHisPheIleArgC 657
     |   :::|||:::::   |||:::::|||:::|||||||:::|||||||
 841 GGAGCAGCTTCTGCAGGTCCTACACAGCACCACGCCCCACTACATTCGCT 890

 657 ysIleLysProAsnSerLysGlnLeuProArgValTyrGluGluAspLeu 673
     ||||||||||||||||||:::          :::::::::   ::::::
 891 GCATCAAGCCCAACAGCCAGGGCCAGGCGCAGACCTTTCTCCAAGAGGAG 940

 674 ValLeuGlnGlnLeuArgCysCysGlyValLeuGluValValArgIleSe 690
     ||||||:::|||||||:::::|||||||:::::|||:::::::::||||||
 941 GTCCTGAGCCAGCTGGAGGCCTGTGGCCTCGTGGAGACCATCCATATCAG 990

 690 rArgSerGlyTyrProThrArgLeuThrHisGlnGluPheAlaGlyArgT 707
     |   :::|||:::|||   |||:::::|||:::::|||:::   ||||
 991 TGCTGCTGGCTTCCCCATCCGGGTCTCTCACCGAAACTTTGTAGAACGAT 1040

 707 yr............GlyPheLeuLeuSer.............AspLys 714
     ||            ||||||:::|||:::              |||:::
1041 ACAAGTTACTAAGAAGGCTTCATCCTTGCACATCCTCTGGCCCCGACAGC 1090

 715 ............Lys.................ValAlaGlnAspP 720
                |||                 :::|||:::   |
1091 CCATATCCTGCCAAAGGGCTCCCTGAATGGTGTCCACACAGCGAGGAAGC 1140

 720 ro...LeuSerValSer.................IleAlaValLeuL 729
     ||   |||::::::|||                 :::   ||||||
1141 CACGCTTGAACCTCTCATCCAGGACATTCTCCACACTCTGCCGGTCCTAA 1190

 729 ysGlnTyrAspValHis.....................ProGluMe 737
        |||   ::   |||                      |||
1191 CTCAGGC.AGCAGCCATAACTGGTGACTCGGCTGAGGCCATGCCAGCCCC 1239

 737 tTyrGlnValGlyTyrThrLysLeuTyrLeuArgThrGlyGlnIleGlyI 754
     :::   |||   ||||||:::::::      :::::::::   :
1240 CATGCACTGTGGCAGGACCAAGGTGTTCATGACTGACTCTATGCTGGAGC 1289

 754 lePheGluAspArgArgLysLysValLeuGlnGlyIleValGlyLeu... 769
     :::::|||      |||   :::|||||:::   :::      |||
1290 TTCTGGAATGTGGGCGTGCCCGGGTGCTGGAGCA.GTGTGCCCGCTGCAT 1338

 770 .GlnLysHisPheArgGlyHisLeuSerArgAlaTyrPheGlnAsnMetA 786
      |||      :::|||   |||       |||         ::::::   |
1339 CCAGGGTGGCTGGAGGCGACACCGGCACCGAGAGCAGGAGCGGCAGTGGC 1388

 786 rgLysValThrLeuValLeuGlnSerTyrIleArgGlyGluAsnAlaArg 802
      ||   |||   |||:::
1389 GGGCCGTCATGCTCATC..............................  1405

 803 ArgLeuPheAspThrGluAlaLysPheHisAlaAspSerValSerGluAl 819

1405 ..............................................  1405

 819 aSerThrAspGluLeuSerAlaValIleHisLeuGlnSerAlaValArgG 836
                                  |||:::|||:::|||:
1406 ..............................CAGGCAGCCATTCGTT 1421

 836 lyTrpLeuAlaArgLysHisPheAsnSerMetGlnArgGlnLysGluLeu 852
     ::|||||::::|||||||||:::                 ::::::|||
1422 CCTGGTTAACTCGGAAACACATC.................CAGAGGCTG 1453

 853 ArgAsnValAlaThrLysSerLysArgLysAlaGlyArgArgIleSerGl 869
     :::   :::|||||||       ||||||
1454 CATGCAGCTGCCACAGTCATCAAGCGT..................... 1480

 869 uAspLysAspIleProLeuGluGlnProGlnValGlnProThrSerMetS 886

1480 ..............................................  1480

 886 erAspLeuGlnLysArgIleLeuLysSerGluAlaAlaLeuSerGlnLys 902

1480 ..............................................  1480

 903 GluGluGluAsnThrAlaLeuArgGluGlnLeuArgGlnPheGluGluAr 919

1481 ................................................GC 1482

 919 gTrpSerGluTyrAspIleLysMet..............LysSerMetG 931
      |||:::::::::   |||:::|||             |||:::::::
1483 ATGGCAGAAGTGGAGAATCAGAATGGCCTGCCTTGCTGCTAAAGAGCTGG 1532

 931 luGluThrTrpGlnLysGlnMet.......SerSerLeu.GlnMetSer. 944
      ::   :::|||:::|||:::         |||||||||  ::::::|||
1533 ATG..GTGTGGAAGAAAAACACTTCTCTCAAGCTCCCTGTTCCCTGAGCA 1580

 945 .LeuAlaAlaAlaArgLysSer.....................LeuAlaA 954
      ||||||||||||   ::::::                     ||||||
1581 CCTCGCCGCTGCAGACCAGGCTCCTGGAGGCAATAATCCGCCTCTGGCCC 1630

 954 laGluSerIleThrGlyGlnAlaGly....Gly..ArgGln...AspThr 967
         :::      ||||||   |||    |||  ||||||   :::
1631 CTGGGACTGGTCCTGGCCAATACGGCTATGGGTGTAGGCAGCTTTCAGAG 1680

 968 SerIleSer.....................ProPheGlyTyrAspSerG 977
     :::|||||||                  |||  |||      |||:
1681 GAAATTAGTGGTCTGGGCTTGCCTCCAGCTCCCCAGGGGCAGCCCCAGTA 1730

 977 luAspThrMetSerThrGly..ThrProGly..ValArgThrProThrAs 992
     ::   |||:::::::|||:::   ||||||  |||      |||:::::
1731 GCTACACTGTCCAGACAGCACAAGACCAGGCTGGTGTCACGTCCATCCGA 1780

 992 nLysPheThrAsnGlyAsn.........ThrProGlu...LeuArgIle 1004
     ::::      :::|||:::          |||   |||   |||:::
1781 GCGCTG.CCTCAGGGATCGATAAAGTTTCACTGCAGAAAGTCTCCACTGC 1829

1005 ArgGluLeuAsnGlySerLeuAsn....AlaValAsnHis.......... 1016
          |||:::::::::|||||||    ::::::::::|||
1830 GGTATGCTGACATCTGCCCTGAACCTTCACCCTACAGCATTACAGGCTTT 1879

1017 ............LeuAlaArg..GluPheAspGlnArgArgLeuAsnPhe 1028
                 |||   |||  :::   |||   |||   |||
1880 AATCAGATTCTGCTGGAAAGACACAGGCTGATCCACGTGACCTCTTCTGC 1929
```

```
    1029 AspGluAspAlaArgAlaIleValGluValLysLeuGlyProGln..... 1043
    1930 CTTCACTGGGCTGGGGTGATCCTTG..GTGCCTTTGTTTCCACAAGGCCT 1977

    1044 .............AlaThrProAsnGlyGlnGlnGlnGlnHisProGluA 1056
    1978 TTTCCTGCCCCCTGCCTTGCCA........................... 1999

    1056 spGluPheArgArgLeuLysLeuArgPheGluThrTrpLysLysAspTyr 1072
    2000 ............................AAGACATTTAATCAGCACACA 2020

    1073 LysAlaArgLeuArgAspThrLysAlaArgLeuHis......ArgValAs 1087
    2021 GCTGCCAGACTATT.CCCACAGTGCTCCAAATGCACATGAACAACAGTGA 2069

    1087 pGlyAspLysGlyArgHisArg                            1094
    2070 CGGCTCCAGCCTTCGACCCAGA                           2091

Sequence name: US09830914C.seq:US-09-830-914C-2

Sequence documentation:
; Sequence 2, Application US/09830914C
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMIC, INC.; TANG, Y. Tom;
;  APPLICANT:  CORLEY, Neil C.; GORGONE, Gina A.;
;  APPLICANT:  GUEGLER, Karl J.; BAUGHN, Mariah R.
;  TITLE OF INVENTION: MYOSIN HEAVY CHAIN HOMOLOG
;  FILE REFERENCE: PF-0621 USN
;  CURRENT APPLICATION NUMBER: US/09/830,914C
;  CURRENT FILING DATE:  2002-11-06
;  PRIOR APPLICATION NUMBER: PCT/US99/26177
;  PRIOR FILING DATE: 1999-11-05
;  PRIOR APPLICATION NUMBER: US 60/172,248
;  PRIOR FILING DATE: 1998-11-05
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PERL Program
; SEQ ID NO 2
;   LENGTH: 2109
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 1929760CB1
US-09-830-914C-2

Alignment of: q39157 x US-09-830-914C-2   ..

Alignment segment 1/1: (-)

                  Quality:  416.20                  Escore:       0
          Matching length:     530              Total length:    1046
 Matching Percent Similarity:  73.40  Matching Percent Identity:  43.21
   Total Percent Similarity:  37.19    Total Percent Identity:  21.89
                     Gaps:      22

Alignment:

      10 ValLysIleSerIleAlaLys............ValSerLeuValGlu.. 21
    2011 ATTAAATG.TCTTTGGCAAGGCAGGGGGCAGGAAAAGGCCTTGTGGAAAC 1963

      22 ...AsnThrGluGluHisAsnLysPro............GluSerGluTr 33
    1962 AAAGGCACCAAGGATCACCCCAGCCCAGTGAAGGCAGAAGAGGTCACGTG 1913

      33 pAsnAsnAsnValGluTyr.........PheIleLysLys.......... 43
    1912 GAT.CAGCCTGTGTCTTTCCAGCAGAATCTGATTAAAGCCTGTAATGCTG 1864

      44 .LysLeuArgValTrpCysArg.ValSerAsnGlyGlnTrpGlnLeuGly 59

    1863 TAGGGTGAAGGTTCAGGGCAGATGTCAGCATACCGCAGTGGAGACTTTCT 1814

      60 LysIle.......................GlnSerThrSerAlaAspTh 68
    1813 GCAGTGAAACTTTATCGATCCCTGAGGCAGCGCTCGGATGGACGTGACAC 1764

      68 rSerLeuValMet..LeuSerThrAlaAsnValValLysValSerThrGl 84
    1763 CAGCCTGGTCTTGTGCTGTCTGGA........................... 1740

      84 uGluLeuPheProAlaAsnProAspIleLeuGluGlyValGluAspLeuI 101

    1740 ..................................................... 1740

     101 leGlnLeuSerTyr...................Leu..AsnGluProSer 110
    1739 ..CAGTGTAGCTACTGGGGCTGCCCCTGGGGAGCTGGAGGCAAGCCCAGA 1692

     111 ValLeu.TyrAsnLeuArgValArgTyrLeuGlnAspValIleTyrSerL 127
    1691 CCACTAATTTCCTCTGAAAGCTGCCTAC............ACCCATAGCC 1654

     127 ys.....AlaGlyProVal.................LeuIleAlaValAs 136
    1653 GTATTGGCCAGGACCAGTCCCAGGGGCCAGAGGCGGATTATTGCC..... 1609

     136 nProPheLysAsnValGluIleTyrGlyAsnAspValIleSerAlaTyrG 153

    1609 ..................................................... 1609

     153 lnLysLysValMetAspAlaProHisValTyrAlaValAlaAspAlaAla 169

    1609 ..................................................... 1609

     170 TyrAspGluMetMetArgGluLysAsnGlnSerLeuIleIleSerGlyGl 186
    1608 .......................TCCAGGAGCCTGGTCTGCAGCGGCGA 1583

     186 u....SerGlyAlaGlyLysThrGluThrAlaLysPhe.......AlaMe 199
    1582 GGTGCTCAGGGAACAGGGAGCTTGAGAGAAGTGTTTTTCTTCCACACCAT 1533

     199 tGlnTyrLeuAlaAlaLeuGlyGlyGlySerCysGlyValGluTyrGluI 216
    1532 CCAGCTCTTT.......................................... 1523

     216 leLeuLysThrThrCysIleLeuGluAlaPheGlyAsnAlaLysThrSer 232

    1523 ..................................................... 1523

     233 ArgAsnAlaAsnSerSerArgPheGlyLysLeuIleGluIleHisPheSe 249
    1522 .............AGCAGCAAGGCAGGCC..ATTCTGATTCTCCACTTCT. 1488

     249 rAlaMetGlyLysIle.....CysGlyAlaLysLeuGluThrPheLeuLe 264
    1487 .GCCATGCA.CGCTTGATGACTGTGGCAGC..................... 1460

     264 uGluLysSerArgValValGlnLeuPheAsnGlyGluArgSerTyrHisI 281

    1460 ..................................................... 1460

     281 lePheTyrGluLeuCysAlaGlyAlaSerProIleLeuLysGluArgLeu 297
    1459 ..............TGCAT.GCAGCCTCTGGATGTGTTTCCGAG...... 1431

     298 LysLeuLysThrAlaSerGluTyrThrTyrLeuSerGlnSerAsp..... 312
    1430 ......TTAACCAGGAACGAATGGCTGC..CTGGATGAGCATGACGGCCC 1389

     313 .....Cys..LeuThrIle.....AlaGlyValAspAspAla........ 322
    1388 GCCACTGCCGCTCCTGCTCTCGGTGCCGGTGTCGCCTCCAGCCACCCTGG 1339
```

```
 323 ...GlnLysPheHisLysLeuLeuGluAlaPheAspIleValGlnIlePr 338
1338 ATGCAGCGGG..CACA..CTGCTCCA.GCACCCGGGCACGCCCACATTCC 1294

 338 oLys....GluHisGlnGluArgAlaPheAlaLeuLeu.....AlaAlaV 352
1293 AGAAGCTCCAGCATA..GAGTCAGTCATGAACACCTTGGTCCTGCCACAG 1246

 352 alLeuTrp..LeuGlyAsnValSerPheArgValThrAsp..... 364
1245 TGCA.TGGGGGCTGGCATGGCCTCAGC.CGAGTCACCAGTTATGGCTGCT 1198

 365 ..........AsnGluAsnHisValGluValValAlaAspGlu..... 375
1197 GCCTGAGTTAGGACCGGCAGAGTGTGGAGAATGTCCTGGATGAGAGGTTC 1148

 376 .AlaValAlaAsnAla.....AlaMetLeu... 383
1147 AAGCGTGGCTTCCTCGCTGTGTGGACACCATTCAGGGAGCCCTTTGGCAG 1098

 384 ...MetGlyCysAsn.....ThrGluGluLeuMetValVa 394
1097 GATATGGGCTGTCGGGGCCAGAGGATGTGCAAGGATGAAGCCTTCTTAGT 1048

 394 lLeuSerThrArgLysLeuGlnAlaGlyThrAspCysIleAla.LysLys 410
1047 A.....ACTTGTATCGTTCTACAAA 1028

 411 LeuThrLeuArg.....GlnAlaThrAspMetArgAs 421
1027 GTTTCGGTGAGAGACCCGGATGGGGAAGCCAGCAGCACTGATATG...GA 981

 421 pGlyIleAlaLysPheIleTyrAlaAsnLeuPhe...AspTrpLeuValG 437
 980 TGGTCTCCACGAGGCCACAGGCCTCCAGCTGGCTCAGGACCTCCTCTTGG 931

 437 luGln...IleAsnIleAlaLeuGluValGlyLysSer..... 448
 930 AGAAAGGTCTGCGCCTGGCCCTGGCTGTTGGGCTTGATGCAGCGAATGTA 881

 449 .....ArgThrGlyArgSerIleSerIleLeuAspI 459
 880 GTGGGGCGTGGTGCTGTGTAGGACCTGCAGAAGCTGCTCC..... 841

 459 leTyrGlyPheGluSerPheLysAsnAsnSerPheGluGlnPheCysIle 475
 841 ..... 841

 476 AsnTyrAlaAsnGluArgLeuGln..GlnHisPheAsnArgHis..... 489
 840 .....AGTGAGGCCTTGAACTTGGACACCACGGTCAACACAGGGGC 800

 490 .....LeuPheLysLeuGluGl 495
 799 CCTGCTCTGGCCAGGGGGTTCCTCCTGGGTCTTCTCTTTGGGGTTAGTAG 750

 495 nGluGluTyrGluGluAspGlyIleAspTrpThrLysValGluPheValA 512
 749 GAAACAGCCCCATGAGCAGGGGGTCC..TGG..... 721

 512 spAsnGlnGluCysLeu...AspLeuIleGluLysLysProIleGlyLeu 527
 720 .....GATTGCTGCAGGAGCCTGGTCAGCTCAGGTGGGATAGGGTC. 680

 528 LeuSerLeuLeuAspGluGluSerAsnPheProLysAlaThrAspLeuTh 544
 679 CTTGTTCTTCTC..... 668

 544 rPheAlaAsnLysLeuLysGlnHisLeuLysThrAsnSerCysPheLysG 561
 668 ..... 668
```

```
 561 lyGluArgGlyArgAlaPheArgValAsnHisTyrAlaGlyGluValLeu 577
 667 .....CACCAGGCCT..GCTGTGTGG 649

 578 TyrAspThrAsn..GlyPheLeuGluLysAsnArgAspProLeuProAla 593
 648 TACCGCACAGGCCCCGCATAATGCACCACAATGAAGCTGGGCTCCCGGCT 599

 594 AspLeuIleAsnLeuLeuSerSerCysAspCysGln.....LeuLeuL 608
 598 GAGCTTATTGTGGCCCAGGCAGGGGCTGCCTGCCAGGGCAGTCTCAATGC 549

 608 ysLeuPheSerThrLysMetArgGlyLysSerGlnLysProLeuMetLeu 624
 548 GTGTCTGGAGCTGGCGT..... 532

 625 SerAspSerThrAsnGlnThrValGlyThrLysPheLysGlyGlnLeuPh 641
 532 ..... 532

 641 eLysLeuMetAsnLysLeuGluAsnThrSerProHisPhe.....IleA 656
 531 .GCGCTGCTGGGTCGATTGAGGCGGC..ATTCCTCATTTATGAGGGAGCA 485

 656 rgCys.....IleLys.ProAsnSerLysGlnLeuPro 666
 484 GATGCTGATGGGGCTTCCCTCAATGAGATCCAAACAGGGCTGGTTGTCCT 435

 667 ArgValTyrGluGlu.AspLeuValLeuGlnGlnLeuArgCys..... 680
 434 G..GTAGTTGATGAATGACC..A..CTCCAGGCCCTCAACTGCGTATTCC 391

 681 ...CysGlyValLeuGluValValArgIleSerArgSerGlyTyrProTh 696
 390 TCCTGCTGGGCCCTTAG.GTAGTG..... 368

 696 rArgLeuThrHisGlnGluPheAlaGlyArgTyrGlyPheLeuLeuSerA 713
 367 .......AGCCAC..... 362

 713 spLysLysValAlaGlnAspProLeuSerValSerIleAlaValLeuLys 729
 362 ..... 362

 730 GlnTyrAspValHisProGluMetTyrGlnValGlyTyrThrLysLeuTy 746
 362 ..... 362

 746 rLeuArgThrGlyGlnIleGlyIlePheGluAspArgArgLysLysValL 763
 361 .....AAAATGCTGC 352

 763 euGln.....GlyIleValGlyLeuGlnLysHisPheArgGlyHis 776
 351 TGCAGCTTCTCATTGGCGTAGTTGATGCACAA..... 320

 777 LeuSerArgAlaTyrPheGlnAsnMetArgLysValThrLeuValLeuGl 793
 320 ..... 320

 793 nSerTyrIleArgGlyGluAsnAlaArgArgLeuPheAspThrGluAlaL 810
 319 .....CTGTTCCAGACTGTTGTCA 301

 810 ys......PheHisAlaAspSerValSerGluAlaSerThrAspGluLeu 824
 300 GGAAATGATTCAAATCC..... 284

 825 SerAlaValIleHisLeuGlnSerAlaValArgGlyTrpLeuAlaArgLy 841
 283 .....ATACACATCCAGCAGGCCTGGGAAGA.TGG..... 255

 841 sHisPheAsnSerMetGlnArg.....GlnLysGluLe 852
```

```
                  ||||||                    ::::::::::
  254 ..............CAGAGAACCCATGGGGCCACTGCAGGGAGCAGAT 221

  852 uArgAsnValAlaThrLysSerLysArgLysAlaGlyArgArgIleSerG 869
      :   |||   :::||||||:::   |||   ||||||
  220 GGGGAATAG.GGGACCAAGGAACCCAGGGTGGCTGGG.............. 185

  869 luAspLysAspIleProLeuGluGlnProGlnValGlnProThrSerMet 885
                 ||||||   :::|||
  184 ..............CCCCTGCCAATGCCTTAT.................. 167

  886 SerAspLeuGlnLysArgIleLeuLysSerGluAlaAlaLeuSerGlnLy 902

  167 ................................................... 167

  902 sGluGluGluAsnThrAlaLeuArgGluGlnLeuArgGlnPheGluGluA 919

  167 ................................................... 167

  919 rgTrpSerGluTyrAspIleLysMetLysSerMetGluGluThrTrpGln 935
                                                 |||||||||
  166 ....................................ACTTGGCAG 158

  936 LysGlnMetSerSerLeuGlnMetSerLeuAlaAlaAlaArgLysSerLe 952

  157 CCT................................................ 155

  952 uAlaAlaGluSerIleThrGlyGlnAlaGlyGlyArgGlnAspThrSerI 969

  155 ................................................... 155

  969 leSerProPheGlyTyrAspSerGluAspThrMetSerThrGlyThrPro 985
                ||||||||||||||   :::|||         ||||||
  154 ...........GGGCATCAGGCTGAGTG.AGCATG.........ACACCT 126

  986 GlyVal...ArgThr..ProThrAsnLysPheThrAsnGly......... 997
        |||   ||||||  |||:::         |||:::|||
  125 GTGGTGGGCAGAACATACCTTCCCCACCCTCCACCCAAGGTGGCTCCCAT 76

  998 .AsnThrPro..........GluLeuArgIleArgGluLeuAsnGlySer 1010
       :::||||||          |||||||||
   75 CTCTACCCCATTCCCCTGTAGAGCTCAGA..................... 47

 1011 LeuAsnAlaValAsnHisLeuAlaArgGluPheAspGlnArgArgLeuAs 1027

   47 ................................................... 47

 1027 nPheAspGluAspAlaArgAlaIleValGluValLysLeuGlyProGlnA 1044
                                                       |
   46 ...................................................G 46

 1044 laThr.......ProAsnGlyGlnGlnGlnGlnHisProGlu         1055
      ||:::       ||||||   :::   ::::::   ||||||
   45 CAAGCAGGAAACCCAAATG..AAGTGTGAAGACAGCCCAGAG         6
```

OM of: US-09-830-914C-2 to: q39157.pep:* out_format : pfs

Date: Wed Apr 2 14:17:30 2003

About: Results were produced by the GenCore software, version 5.1.4_p5_4578,
Copyright (c) 1993-2003 Compugen Ltd.

Command line parameters:
-MODEL=frame+_n2p.model -DEV=soft -Q=US09830914C.seq -DB=q39157.pep
-SUFFIX=pfs -OUT=compare2 -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits
-START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY
-NEG_SCORES=0 -LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-FGAPEXT=0.1 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=0.1

Search information:
Query: US-09-830-914C-2
Query length: 2109
Database: q39157.pep:*
Database sequences: 1
Database length: 1101
Search time (sec): 1

Score list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation ..
q39157.pep:q39157	+	850.51	0.00	0.0	1101	TOIG of: q39157 check: 9498
q39157.pep:q39157	-	416.20	0.00	0.0	1101	TOIG of: q39157 check: 9498

Sequence name: q39157.pep:q39157

Sequence documentation:
; TOIG of: q39157 check: 9498 from: 1 to: 1101
;
; ID Q39157 PRELIMINARY; PRT; 1101 AA.
; AC Q39157;
; DT 01-NOV-1996 (TrEMBLrel. 01, Created)
; DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
; DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
; DE Myosin (Fragment).
; GN MYOSIN HEAVY CHAIN.
; OS Arabidopsis thaliana (Mouse-ear cress).
; OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
; OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
; OX NCBI_TaxID=3702;
; RN [1]
; RP SEQUENCE FROM N.A.
; RC STRAIN=COLUMBIA ECOTYPE;
; RX MEDLINE=95111097; PubMed=7811972;
; RA Kinkema M.D., Wang H., Schiefelbein J.;
; RT "Molecular analysis of the myosin gene family of Arabidopsis
; RT thaliana.";
; RL Plant Mol. Biol. 26:1139-1153(1994).
; DR EMBL; Z34292; CAA84065.1; -.
; DR HSSP; P08799; 1MND.
; DR InterPro; IPR000048; IQ_region.
; DR InterPro; IPR001609; myosin_head.
; DR InterPro; IPR004009; Myosin_N.
; DR Pfam; PF00612; IQ; 3.
; DR Pfam; PF00063; myosin_head; 1.
; DR Pfam; PF02736; Myosin_N; 1.
; DR PRINTS; PR00193; MYOSINHEAVY.
; DR ProDom; PD000355; myosin_head; 1.
; DR SMART; SM00015; IQ; 3.
; DR SMART; SM00242; MYSc; 1.
; DR PROSITE; PS50096; IQ; 1.
; FT NON_TER 1 1
; SQ SEQUENCE 1101 AA; 125695 MW; F9BF30291A26A184 CRC64;
;
; Q39157 Length: 1101 April 2, 2003 13:55 Type: P Check: 9498 ..
q39157

Alignment of: US-09-830-914C-2 x q39157 ..

Alignment segment 1/1: (+)

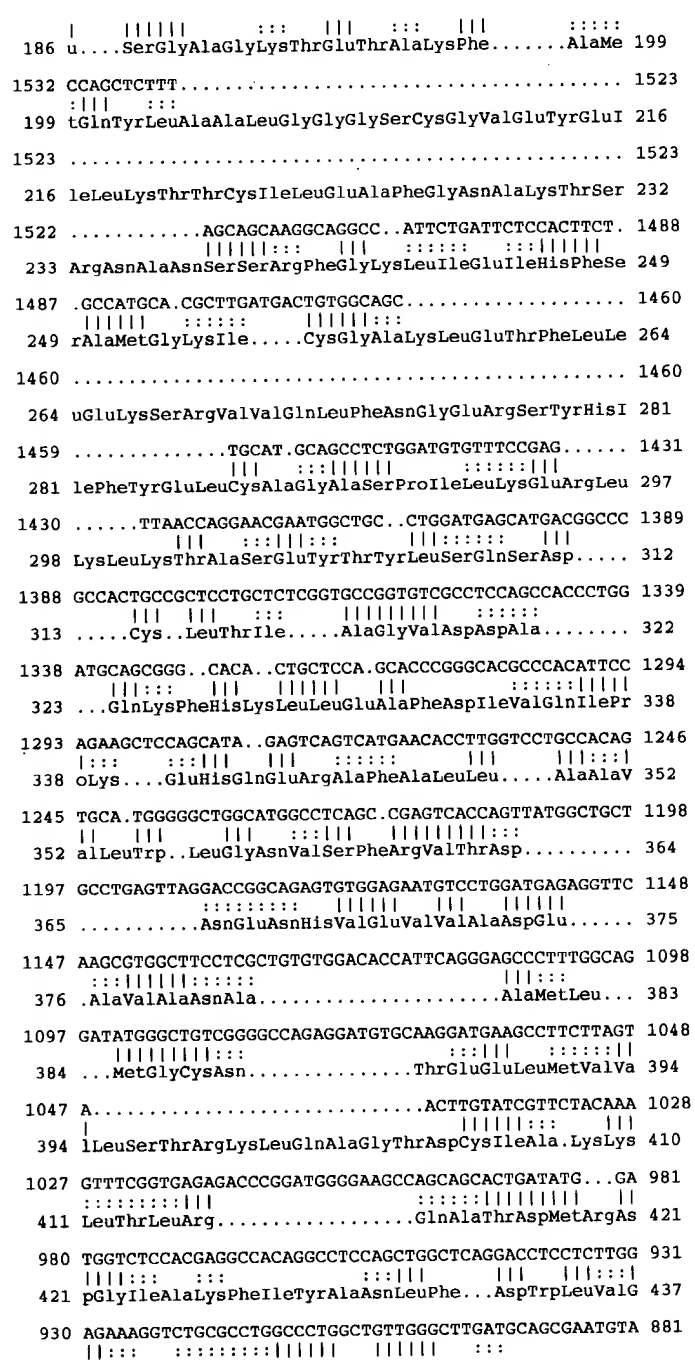
Quality:	850.51	Escore:	0
Matching length:	577	Total length:	885
Matching Percent Similarity:	75.39	Matching Percent Identity:	44.37
Total Percent Similarity:	49.15	Total Percent Identity:	28.93
Gaps:	11		

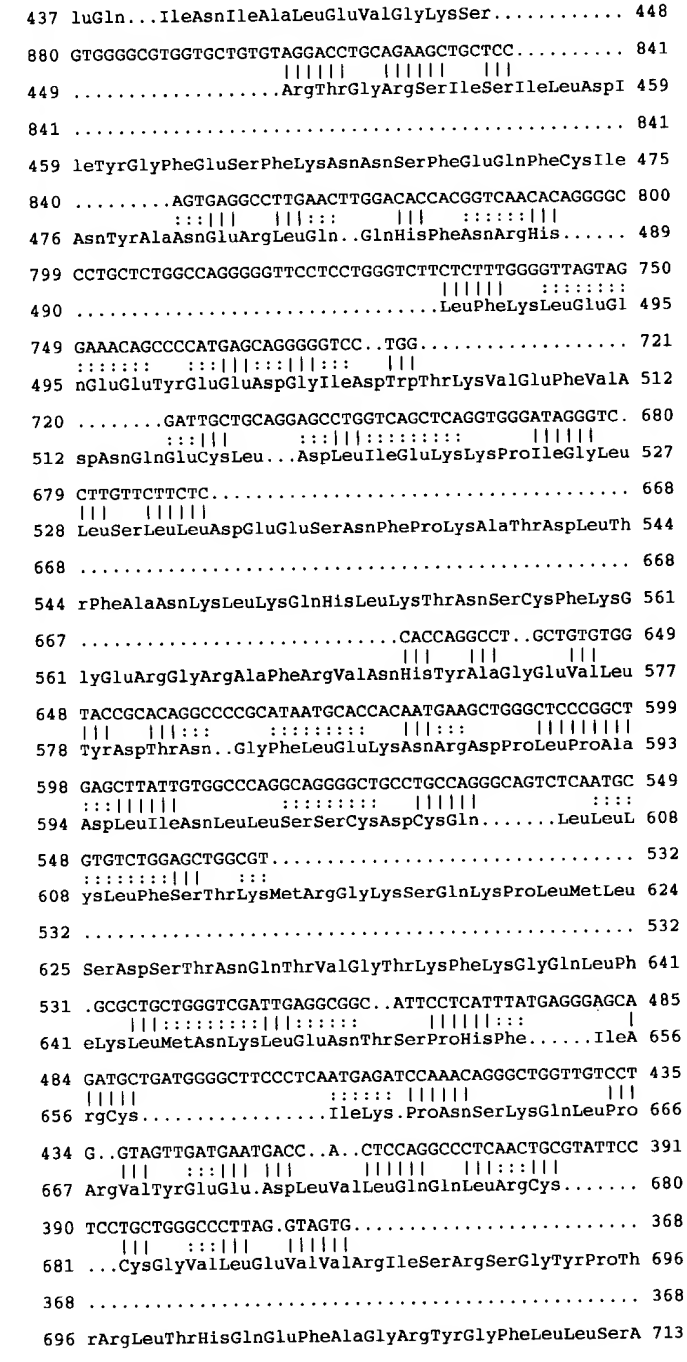
Alignment:

```
      4 TGCTCTGGGCTGTCTTCACACTTCATTTGGGTTTCCTGCTTGCTCTGAGC 53
        |||   |||::::::   :::::::::   ::::::|||:::||| |||
    210 Cys...GlyValGluTyrGluIleLeuLysThrThrCysIleLeu.GluA 225

     54 TCTACAGGGGAATGGG..................................... 69
          :::  ||||||:::
    225 laPhe..GlyAsnAlaLysThrSerArgAsnAlaAsnSerSerArgPheG 241

     70 ........GTAGAGATGGGAGCCACCTTGGGTGGAGGGTGGGGAAGGTAT 111
                ::::|||:::     |||::::::      :::||||||:::|
    241 lyLysLeuIleGluIle.....HisPheSer...AlaMetGlyLysIleC 255

    112 GTTCTGCC.......................................... 119
        ||:::|||
    255 ysGlyAlaLysLeuGluThrPheLeuLeuGluLysSerArgValValGln 271

    120 ................................CACCACAGGTGTCATGC 136
                                        ::::::   |||   ::
    272 LeuPheAsnGlyGluArgSerTyrHisIlePheTyrGluLeuCysAlaGl 288

    137 TCACTCAGCCT....................................... 147
        :   ||| |||
    288 yAlaSer.ProIleLeuLysGluArgLeuLysLeuLysThrAlaSerGlu 304

    148 ..............................................GA 149
                                                      ||
    305 TyrThrTyrLeuSerGlnSerAspCysLeuThrIleAlaGlyValAspAs 321

    150 TGCCCAGGCTGCCAAGTATAAGGCATTGGCAGGGGCCCAGCCACCCTGGG 199
        |||||||      |||:::      |||      |||
    321 pAlaGln......LysPheHisLysLeuLeuGluAla.............. 331

    200 TTCC..TTGGTCCCCTATTCCCCAT.......................C 223
        |||   :::|||    ||||||                         |
    332 PheAspIleVal.GlnIleProLysGluHisGlnGluArgAlaPheAlaL 348

    224 TGCTCCCTGCAGTGGCCCCATGGGTTCTCTGCCATCTTCCCAGGC..... 268
        |||||   ||||||      |||:::                 |||
    348 euLeuAlaAlaValLeu...TrpLeu..............GlyAsnVa 358

    268 ..................................................... 268

    358 lSerPheArgValThrAspAsnGluAsnHisValGluValValAlaAspG 375

    268 ..................................................... 268

    375 luAlaValAlaAsnAlaAlaMetLeuMetGlyCysAsnThrGluGluLeu 391

    268 ..................................................... 268

    392 MetValValLeuSerThrArgLysLeuGlnAlaGlyThrAspCysIleAl 408

    268 ..................................................... 268

    408 aLysLysLeuThrLeuArgGlnAlaThrAspMetArgAspGlyIleAlaL 425

    268 ..................................................... 268

    425 ysPheIleTyrAlaAsnLeuPheAspTrpLeuValGluGlnIleAsnIle 441

    269 ..................................CTGCTGGA 276
```

```
                                                      :::|||||
  442 AlaLeuGluValGlyLysSerArgThrGlyArgSerIleSerIleLeuAs 458

  277 TGTGTATGGATTTGAATCATTTCCTGACAACAGTCTGGAACAGTTGTGCA 326
      |:::||||||||||||||||||   :::|||||||:::|||||||:::||||
  458 pIleTyrGlyPheGluSerPheLysAsnAsnSerPheGluGlnPheCysI 475

  327 TCAACTACGCCAATGAGAAGCTGCAGCAGCATTTTGTGGCTCACTACCTA 376
      |||||||||||||||||:::|||||||||||||||||      |||   :::
  475 leAsnTyrAlaAsnGluArgLeuGlnGlnHisPheAsnArgHisLeuPhe 491

  377 AGGGCCCAGCAGGAGGAATACGCAGTTGAGGGCCTGGAGTGGTCATTCAT 426
      :::   :::||||||||||||      :::|||:::::|||:::   ::
  492 LysLeuGluGlnGluGluTyrGluGluAspGlyIleAspTrpThrLysVa 508

  427 CAACTACCAGGACAACCAGCCCTGTTTGGATCTCATTGAGGGAAGCCCCA 476
      :::::::   |||||||||   ||||||||||||||||||   :::||||
  508 lGluPheValAspAsnGlnGluCysLeuAspLeuIleGluLysLysProI 525

  477 TCAGCATCTGCTCCCTCATAAATGAGGAATGCCGCCTCAATCGACCCAGC 526
      ||:::::::   ||||||::::::||||||         |||   |||:::
  525 leGlyLeuLeuSerLeuLeuAspGluGluSer......AsnPheProLys 539

  527 AGC..............GCACGCCAGCTCCAGACACGCATTGAGACTGC 561
      :::               |||::::::|||:::   :::::::::|||
  540 AlaThrAspLeuThrPheAlaAsnLysLeuLysGlnHisLeuLysThr.. 555

  562 CCTGGCAGGCAGCCCCTGCCTGGGCCACAATAAGCTCAGCCGGGAGCCCA 611
              :::   |||:::          |||   :::|||      :
  556 ..........AsnSerCysPhe.........LysGlyGluArgGlyArgA 566

  612 GCTTCATTGTGGTGCATTATGCGGGGCCTGTGCGGTACCACACAGCAGGC 661
      ::|||   |||   ||||||||||||   |||   |||   |||   |||
  566 laPheArgValAsnHisTyrAlaGlyGluValLeuTyrAspThrAsnGly 582

  662 CTGGTGGAGAAGAACAAGGACCCTATCCCACCTGAGCTGACCAGGCTCCT 711
      ::::::|||||||||:::|||||||:::|||   :::|||   :::|||||
  583 PheLeuGluLysAsnArgAspProLeuProAlaAspLeuIleAsnLeuLe 599

  712 GCAGCAATCCCAGGACCCCCTGCTCATGGGGCTGTTTCCTACTAACCCCA 761
      |   :::|||   |||      |||:::   ||||||   |||
  599 u...SerSerCysAspCysGlnLeuLeuLysLeuPheSerThr....... 612

  762 AAGAGAAGACCCAGGAGGAACCCCCTGGCCAGAGCAGGGCCCCTGTGTTG 811
      |||   |||                 |||:::|||:::   |||::::::
  613 LysMetArg.................GlyLysSerGlnLysProLeuMet 623

  812 .....................ACCGTGGTGTCCAAGTTCAAGGCCTCACT 840
                            ||||||   :::|||||||||:::::||
  624 LeuSerAspSerThrAsnGlnThrValGlyThrLysPheLysGlyGlnLe 640

  841 GGAGCAGCTTCTGCAGGTCCTACACAGCACCACGCCCCCACTACATTCGCT 890
      |   :::|||::::::   |||:::::::|||:::|||||||:::|||||||
  640 uPheLysLeuMetAsnLysLeuGluAsnThrSerProHisPheIleArgC 657

  891 GCATCAAGCCCAACAGCCAGGGCCAGGCGCAGACCTTTCTCCAAGAGGAG 940
      |||||||||||||||||:::         :::::::::   ::::::
  657 ysIleLysProAsnSerLysGlnLeuProArgValTyrGluGluAspLeu 673

  941 GTCCTGAGCCAGCTGGAGGCCTGTGGCCTCGTGGAGACCATCCATATCAG 990
      ||||||:::|||||:::::|||||||:::::||||:::::::::|||||
  674 ValLeuGlnGlnLeuArgCysCysGlyValLeuGluValValArgIleSe 690

  991 TGCTGCTGGCTTCCCCATCCGGGTCTCTCACCGAAACTTTGTAGAACGAT 1040
      |   :::|||:::|||   |||:::::::|||:::::|||:::   ||||
  690 rArgSerGlyTyrProThrArgLeuThrHisGlnGluPheAlaGlyArgT 707

 1041 ACAAGTTACTAAGAAGGCTTCATCCTTGCACATCCTCTGGCCCCGACAGC 1090
      ||              ||||||:::|||:::              |||:::
  707 yr............GlyPheLeuLeuSer.............AspLys 714

 1091 CCATATCCTGCCAAAGGGCTCCCTGAATGGTGTCCACACAGCGAGGAAGC 1140
                                            :::|||:::   |

  715 ...............................LysValAlaGlnAspP 720

 1141 CACGCTTGAACCTCTCATCCAGGACATTCTCCACACTCTGCCCGGTCCTAA 1190
      ||   |||::::::|||                    :::   ||||||
  720 ro...LeuSerValSer.................IleAlaValLeuL 729

 1191 CTCAGGC.AGCAGCCATAACTGGTGACTCGGCTGAGGCCATGCCAGCCCC 1239
        |||   :::   |||                          |||
  729 ysGlnTyrAspValHis.......................ProGluMe 737

 1240 CATGCACTGTGGCAGGACCAAGGTGTTCATGACTGACTCTATGCTGGAGC 1289
         :::   |||   ||||||::::::::      :::::::::   :
  737 tTyrGlnValGlyTyrThrLysLeuTyrLeuArgThrGlyGlnIleGlyI 754

 1290 TTCTGGAATGTGGGCGTGCCCGGGTGCTGGAGCA.GTGTGCCCGCTGCAT 1338
      :::::|||      |||   :::||||||:::   :::      |||
  754 lePheGluAspArgArgLysLysValLeuGlnGlyIleValGlyLeu... 769

 1339 CCAGGGTGGCTGGAGGCGACACCGGCACCGAGAGCAGGAGCGGCAGTGGC 1388
       |||      :::|||   |||      |||          ::::::   |
  770 .GlnLysHisPheArgGlyHisLeuSerArgAlaTyrPheGlnAsnMetA 786

 1389 GGGCCGTCATGCTCATC............................... 1405
      ||   |||   |||:::
  786 rgLysValThrLeuValLeuGlnSerTyrIleArgGlyGluAsnAlaArg 802

 1405 ............................................... 1405

  803 ArgLeuPheAspThrGluAlaLysPheHisAlaAspSerValSerGluAl 819

 1406 ...............................CAGGCAGCCATTCGTT 1421
                                     |||:::|||:::|||:
  819 aSerThrAspGluLeuSerAlaValIleHisLeuGlnSerAlaValArgG 836

 1422 CCTGGTTAACTCGGAAACACATC.................CAGAGGCTG 1453
      ::|||||||:::|||||||||:::                 :::::::|||
  836 lyTrpLeuAlaArgLysHisPheAsnSerMetGlnArgGlnLysGluLeu 852

 1454 CATGCAGCTGCCACAGTCATCAAGCGT..................... 1480
      :::   :::||||||      ||||||
  853 ArgAsnValAlaThrLysSerLysArgLysAlaGlyArgArgIleSerGl 869

 1480 ............................................... 1480

  869 uAspLysAspIleProLeuGluGlnProGlnValGlnProThrSerMetS 886

 1480 ............................................... 1480

  886 erAspLeuGlnLysArgIleLeuLysSerGluAlaAlaLeuSerGlnLys 902

 1481 .............................................GC 1482

  903 GluGluGluAsnThrAlaLeuArgGluGlnLeuArgGlnPheGluGluAr 919

 1483 ATGGCAGAAGTGGAGAATCAGAATGGCCTGCCTTGCTGCTAAAGAGCTGG 1532
      |||:::::::::   |||:::|||               |||::::::::
  919 gTrpSerGluTyrAspIleLysMet..............LysSerMetG 931

 1533 ATG..GTGTGGAAGAAAAACACTTCTCTCAAGCTCCCTGTTCCCTGAGCA 1580
      ::   :::|||:::|||:::          |||||||| :::::|||
  931 luGluThrTrpGlnLysGln.......MetSerSerLeu.GlnMetSer. 944

 1581 CCTCGCCGCTGCAGACCAGGCTCCTGGAGGCAATAATCCGCCTCTGGCCC 1630
       ||||||||||||   ::::::                  ||||||
  945 .LeuAlaAlaAlaArgLysSer...................LeuAlaA 954

 1631 CTGGGACTGGTCCTGGCCAATACGGCTATGGGTGTAGGCAGCTTTCAGAG 1680
         :::      ||||||   |||    |||  ||||||   :::
  954 laGluSerIleThrGlyGlnAlaGly....Gly..ArgGln...AspThr 967

 1681 GAAATTAGTGGTCTGGGCTTGCCTCCAGCTCCCCAGGGGCAGCCCCAGTA 1730
      :::||||||                 |||   |||     |||:
  968 SerIleSer.....................ProPheGlyTyrAspSerG 977
```

```
 1731 GCTACACTGTCCAGACAGCACAAGACCAGGCTGGTGTCACGTCCATCCGA 1780
      ::   |||::::::|||:::     ||||||  |||      |||:::::
  977 luAspThrMetSerThrGly..ThrProGly..ValArgThrProThrAs 992

 1781 GCGCTG.CCTCAGGGATCGATAAAGTTTCACTGCAGAAAGTCTCCACTGC 1829
      ::::      :::|||:::          |||   |||   |||:::
  992 nLysPheThrAsnGlyAsn.....ThrProGlu...LeuArgIle 1004

 1830 GGTATGCTGACATCTGCCCTGAACCTTCACCCTACAGCATTACAGGCTTT 1879
            |||::::::::::||||||    ::::::::::|||
 1005 ArgGluLeuAsnGlySerLeuAsn....AlaValAsnHis..... 1016

 1880 AATCAGATTCTGCTGGAAAGACACAGGCTGATCCACGTGACCTCTTCTGC 1929
                 |||   |||  :::   |||   |||   |||
 1017 ............LeuAlaArg..GluPheAspGlnArgArgLeuAsnPhe 1028

 1930 CTTCACTGGGCTGGGGTGATCCTTG..GTGCCTTTGTTTCCACAAGGCCT 1977
        :::   |||   :::|||:::   |||   |||   ||||||
 1029 AspGluAspAlaArgAlaIleValGluValLysLeuGlyProGln..... 1043

 1978 TTTCCTGCCCCCTGCCTTGCCA..... 1999
                   |||   |||
 1044 .............AlaThrProAsnGlyGlnGlnGlnGlnHisProGluA 1056

 2000 ............................AAGACATTTAATCAGCACACA 2020
                                  :::|||::::::::
 1056 spGluPheArgArgLeuLysLeuArgPheGluThrTrpLysLysAspTyr 1072

 2021 GCTGCCAGACTATT.CCCACAGTGCTCCAAATGCACATGAACAACAGTGA 2069
        |||||||||       |||      ::::::|||       :::    ||
 1073 LysAlaArgLeuArgAspThrLysAlaArgLeuHis......ArgValAs 1087

 2070 CGGCTCCAGCCTTCGACCCAGA                              2091
      ||||::::::   |||   |||
 1087 pGlyAspLysGlyArgHisArg                              1094

Sequence name: q39157.pep:q39157

Sequence documentation:
;  TOIG of: q39157  check: 9498  from: 1  to: 1101
;
; ID   Q39157      PRELIMINARY;      PRT;  1101 AA.
; AC   Q39157;
; DT   01-NOV-1996 (TrEMBLrel. 01, Created)
; DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
; DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
; DE   Myosin (Fragment).
; GN   MYOSIN HEAVY CHAIN.
; OS   Arabidopsis thaliana (Mouse-ear cress).
; OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
; OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
; OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
; OX   NCBI_TaxID=3702;
; RN   [1]
; RP   SEQUENCE FROM N.A.
; RC   STRAIN=COLUMBIA ECOTYPE;
; RX   MEDLINE=95111097; PubMed=7811972;
; RA   Kinkema M.D., Wang H., Schiefelbein J.;
; RT   "Molecular analysis of the myosin gene family of Arabidopsis
; RT   thaliana.";
; RL   Plant Mol. Biol. 26:1139-1153(1994).
; DR   EMBL; Z34292; CAA84065.1; -.
; DR   HSSP; P08799; 1MND.
; DR   InterPro; IPR000048; IQ_region.
; DR   InterPro; IPR001609; myosin_head.
; DR   InterPro; IPR004009; Myosin_N.
; DR   Pfam; PF00612; IQ; 3.
; DR   Pfam; PF00063; myosin_head; 1.
; DR   Pfam; PF02736; Myosin_N; 1.
; DR   PRINTS; PR00193; MYOSINHEAVY.
; DR   ProDom; PD000355; myosin_head; 1.
; DR   SMART; SM00015; IQ; 3.
; DR   SMART; SM00242; MYSc; 1.
; DR   PROSITE; PS50096; IQ; 1.
; FT   NON_TER      1     1
; SQ   SEQUENCE   1101 AA;  125695 MW;  F9BF30291A26A184 CRC64;
;
;   Q39157  Length: 1101  April 2, 2003 13:55  Type: P  Check: 9498  ..
q39157

Alignment of: US-09-830-914C-2 x q39157  ..

Alignment segment 1/1: (-)

                    Quality:  416.20                    Escore:       0
              Matching length:     530              Total length:    1046
 Matching Percent Similarity:   73.21  Matching Percent Identity:   43.02
   Total Percent Similarity:   37.09    Total Percent Identity:   21.80
                       Gaps:      21

Alignment:

 2011 ATTAAATG.TCTTTGGCAAGGCAGGGGGCAGGAAAAGGCCTTGTGGAAAC 1963
      :::|||   |||:::|||:::                :::|||||||||
   10 ValLysIleSerIleAlaLys.............ValSerLeuValGlu.. 21

 1962 AAAGGCACCAAGGATCACCCCAGCCCAGTGAAGGCAGAAGAGGTCACGTG 1913
        :::|||::::::|||   :::|||           |||      ||
   22 ...AsnThrGluGluHisAsnLysPro............GluSerGluTr 33

 1912 GAT.CAGCCTGTGTCTTTCCAGCAGAATCTGATTAAAGCCTGTAATGCTG 1864
      |   :::   |||::::::         :::||||||                 '
   33 pAsnAsnAsnValGluTyr.........PheIleLysLys.......... 43

 1863 TAGGGTGAAGGTTCAGGGCAGATGTCAGCATACCGCAGTGGAGACTTTCT 1814
       ::::::::::|||      ||| ||||||      ||||||:::|||:::
   44 .LysLeuArgValTrpCysArg.ValSerAsnGlyGlnTrpGlnLeuGly 59

 1813 GCAGTGAAACTTTATCGATCCCTGAGGCAGCGCTCGGATGGACGTGACAC 1764
         :::                        :::|||   :::   |||||
   60 LysIle.....................GlnSerThrSerAlaAspTh 68

 1763 CAGCCTGGTCTTGTGCTGTCTGGA..... 1740
      ||||||||||::  ||||||
   68 rSerLeuValMet..LeuSerThrAlaAsnValValLysValSerThrGl 84

 1740 ..... 1740

   84 uGluLeuProAlaAsnProAspIleLeuGluGlyValGluAspLeuI 101

 1739 ..CAGTGTAGCTACTGGGGCTGCCCCTGGGGAGCTGGAGGCAAGCCCAGA 1692
        |||   ||||||                 |||  :::::::|||
  101 leGlnLeuSerTyr.................Leu..AsnGluProSer 110

 1691 CCACTAATTTCCTCTGAAAGCTGCCTAC...........ACCCATAGCC 1654
        ||| :::   |||::::::   |||              :::|||:
  111 ValLeu.TyrAsnLeuArgValArgTyrLeuGlnAspValIleTyrSerL 127

 1653 GTATTGGCCAGGACCAGTCCCAGGGGCCAGAGGCGGATTATTGCC..... 1609
      ::      ||||||||               :::||||||
  127 ys.....AlaGlyProVal..............LeuIleAlaValAs 136

 1609 ..... 1609

  136 nProPheLysAsnValGluIleTyrGlyAsnAspValIleSerAlaTyrG 153

 1609 ..... 1609

  153 lnLysLysValMetAspAlaProHisValTyrAlaValAlaAspAlaAla 169

 1608 ....................TCCAGGAGCCTGGTCTGCAGCGGCGA 1583
                          ::::::||||||:::   ||||||||
  170 TyrAspGluMetMetArgGluLysAsnGlnSerLeuIleIleSerGlyGl 186

 1582 GGTGCTCAGGGAACAGGGAGCTTGAGAGAAGTGTTTTTCTTCCACACCAT 1533
```


186 u....SerGlyAlaGlyLysThrGluThrAlaLysPhe.....AlaMe 199
1532 CCAGCTCTTT..... 1523
199 tGlnTyrLeuAlaAlaLeuGlyGlyGlySerCysGlyValGluTyrGluI 216
1523 1523
216 leLeuLysThrThrCysIleLeuGluAlaPheGlyAsnAlaLysThrSer 232
1522 .....AGCAGCAAGGCAGGCC..ATTCTGATTCTCCACTTCT. 1488
233 ArgAsnAlaAsnSerSerArgPheGlyLysLeuIleGluIleHisPheSe 249
1487 .GCCATGCA.CGCTTGATGACTGTGGCAGC..... 1460
249 rAlaMetGlyLysIle.....CysGlyAlaLysLeuGluThrPheLeuLe 264
1460 1460
264 uGluLysSerArgValValGlnLeuPheAsnGlyGluArgSerTyrHisI 281
1459 .....TGCAT.GCAGCCTCTGGATGTGTTTCCGAG..... 1431
281 lePheTyrGluLeuCysAlaGlyAlaSerProIleLeuLysGluArgLeu 297
1430 ......TTAACCAGGAACGAATGGCTGC..CTGGATGAGCATGACGGCCC 1389
298 LysLeuLysThrAlaSerGluTyrThrTyrLeuSerGlnSerAsp..... 312
1388 GCCACTGCCGCTCCTGCTCTCGGTGCCGGTGTCGCCTCCAGCCACCCTGG 1339
313 .....Cys..LeuThrIle.....AlaGlyValAspAspAla..... 322
1338 ATGCAGCGGG..CACA..CTGCTCCA.GCACCCGGGCACGCCCACATTCC 1294
323 ...GlnLysPheHisLysLeuLeuGluAlaPheAspIleValGlnIlePr 338
1293 AGAAGCTCCAGCATA..GAGTCAGTCATGAACACCTTGGTCCTGCCACAG 1246
338 oLys....GluHisGlnGluArgAlaPheAlaLeuLeu.....AlaAlaV 352
1245 TGCA.TGGGGGCTGGCATGGCCTCAGC.CGAGTCACCAGTTATGGCTGCT 1198
352 alLeuTrp..LeuGlyAsnValSerPheArgValThrAsp..... 364
1197 GCCTGAGTTAGGACCGGCAGAGTGTGGAGAATGTCCTGGATGAGAGGTTC 1148
365 .....AsnGluAsnHisValGluValValAlaAspGlu..... 375
1147 AAGCGTGGCTTCCTCGCTGTGTGGACACCATTCAGGGAGCCCTTTGGCAG 1098
376 .AlaValAlaAsnAla.....AlaMetLeu... 383
1097 GATATGGGCTGTCGGGGCCAGAGGATGTGCAAGGATGAAGCCTTCTTAGT 1048
384 ...MetGlyCysAsn.....ThrGluGluLeuMetValVa 394
1047 A.....ACTTGTATCGTTCTACAAA 1028
394 lLeuSerThrArgLysLeuGlnAlaGlyThrAspCysIleAla.LysLys 410
1027 GTTTCGGTGAGAGACCCGGATGGGGAAGCCAGCAGCACTGATATG...GA 981
411 LeuThrLeuArg.....GlnAlaThrAspMetArgAs 421
980 TGGTCTCCACGAGGCCACAGGCCTCCAGCTGGCTCAGGACCTCCTCTTGG 931
421 pGlyIleAlaLysPheIleTyrAlaAsnLeuPhe...AspTrpLeuValG 437
930 AGAAAGGTCTGCGCCTGGCCCTGGCTGTTGGGCTTGATGCAGCGAATGTA 881


437 luGln...IleAsnIleAlaLeuGluValGlyLysSer..... 448
880 GTGGGGCGTGGTGCTGTGTAGGACCTGCAGAAGCTGCTCC..... 841
449 .....ArgThrGlyArgSerIleSerIleLeuAspI 459
841 841
459 leTyrGlyPheGluSerPheLysAsnAsnSerPheGluGlnPheCysIle 475
840 .....AGTGAGGCCTTGAACTTGGACACCACGGTCAACACAGGGGC 800
476 AsnTyrAlaAsnGluArgLeuGln..GlnHisPheAsnArgHis..... 489
799 CCTGCTCTGGCCAGGGGGTTCCTCCTGGGTCTTCTCTTTGGGGTTAGTAG 750
490 .....LeuPheLysLeuGluGl 495
749 GAAACAGCCCCATGAGCAGGGGGTCC..TGG..... 721
495 nGluGluTyrGluGluAspGlyIleAspTrpThrLysValGluPheValA 512
720 .....GATTGCTGCAGGAGCCTGGTCAGCTCAGGTGGGATAGGGTC. 680
512 spAsnGlnGluCysLeu...AspLeuIleGluLysLysProIleGlyLeu 527
679 CTTGTTCTTCTC..... 668
528 LeuSerLeuLeuAspGluGluSerAsnPheProLysAlaThrAspLeuTh 544
668 668
544 rPheAlaAsnLysLeuLysGlnHisLeuLysThrAsnSerCysPheLysG 561
667 .....CACCAGGCCT..GCTGTGTGG 649
561 lyGluArgGlyArgAlaPheArgValAsnHisTyrAlaGlyGluValLeu 577
648 TACCGCACAGGCCCCGCATAATGCACCACAATGAAGCTGGGCTCCCGGCT 599
578 TyrAspThrAsn..GlyPheLeuGluLysAsnArgAspProLeuProAla 593
598 GAGCTTATTGTGGCCCAGGCAGGGGCTGCCTGCCAGGGCAGTCTCAATGC 549
594 AspLeuIleAsnLeuLeuSerSerCysAspCysGln.....LeuLeuL 608
548 GTGTCTGGAGCTGGCGT..... 532
608 ysLeuPheSerThrLysMetArgGlyLysSerGlnLysProLeuMetLeu 624
532 532
625 SerAspSerThrAsnGlnThrValGlyThrLysPheLysGlyGlnLeuPh 641
531 .GCGCTGCTGGGTCGATTGAGGCGGC..ATTCCTCATTTATGAGGGAGCA 485
641 eLysLeuMetAsnLysLeuGluAsnThrSerProHisPhe.....IleA 656
484 GATGCTGATGGGGCTTCCCTCAATGAGATCCAAACAGGGCTGGTTGTCCT 435
656 rgCys.....IleLys.ProAsnSerLysGlnLeuPro 666
434 G..GTAGTTGATGAATGACC..A..CTCCAGGCCCTCAACTGCGTATTCC 391
667 ArgValTyrGluGlu.AspLeuValLeuGlnGlnLeuArgCys..... 680
390 TCCTGCTGGGCCCTTAG.GTAGTG..... 368
681 ...CysGlyValLeuGluValValArgIleSerArgSerGlyTyrProTh 696
368 368
696 rArgLeuThrHisGlnGluPheAlaGlyArgTyrGlyPheLeuLeuSerA 713

```
368 ..................................................  368

713 spLysLysValAlaGlnAspProLeuSerValSerIleAlaValLeuLys  729

368 ..................................................  368

730 GlnTyrAspValHisProGluMetTyrGlnValGlyTyrThrLysLeuTy  746

367 ..............................AGCCACAAAATGCTGC  352
                                     :::|||   :::|
746 rLeuArgThrGlyGlnIleGlyIlePheGluAspArgArgLysLysValL  763

351 TGCAGCTTCTCATTGGCGTAGTTGATGCACAA..................  320
    |||||         |||:::|||      |||
763 euGln.........GlyIleValGlyLeuGlnLysHisPheArgGlyHis  776

320 ..................................................  320

777 LeuSerArgAlaTyrPheGlnAsnMetArgLysValThrLeuValLeuGl  793

319 ............................CTGTTCCAGACTGTTGTCA  301
                                 ||||||:::|||   ::::
793 nSerTyrIleArgGlyGluAsnAlaArgArgLeuPheAspThrGluAlaL  810

300 GGAAATGATTCAAATCC..............................  284
    ::      |||   :::
810 ys......PheHisAlaAspSerValSerGluAlaSerThrAspGluLeu  824

283 .........ATACACATCCAGCAGGCCTGGGAAGA.TGG...........  255
            ||||||:::|||:::|||   :::   |||
825 SerAlaValIleHisLeuGlnSerAlaValArgGlyTrpLeuAlaArgLy  841

254 ..............CAGAGAACCCATGGGGCCACTGCAGGGAGCAGAT  221
                    ||||||               ::::::::::
841 sHisPheAsnSerMetGlnArg.............GlnLysGluLe  852

220 GGGGAATAG.GGGACCAAGGAACCCAGGGTGGCTGGG.............  185
    :   |||   :::||||||:::   |||   ||||||
852 uArgAsnValAlaThrLysSerLysArgLysAlaGlyArgArgIleSerG  869

184 ............CCCCTGCCAATGCCTTAT..................  167
                  ||||||   :::|||
869 luAspLysAspIleProLeuGluGlnProGlnValGlnProThrSerMet  885

167 ..................................................  167

886 SerAspLeuGlnLysArgIleLeuLysSerGluAlaAlaLeuSerGlnLy  902

167 ..................................................  167

902 sGluGluGluAsnThrAlaLeuArgGluGlnLeuArgGlnPheGluGluA  919

166 ......................................ACTTGGCAG  158
                                          |||||||||
919 rgTrpSerGluTyrAspIleLysMetLysSerMetGluGluThrTrpGln  935

157 CCT...............................................  155

936 LysGlnMetSerSerLeuGlnMetSerLeuAlaAlaAlaArgLysSerLe  952

155 ..................................................  155

952 uAlaAlaGluSerIleThrGlyGlnAlaGlyGlyArgGlnAspThrSerI  969

154 ..........GGGCATCAGGCTGAGTG.AGCATG.........ACACCT  126
              |||::::::::::|||   :::|||          ||||||
969 leSerProPheGlyTyrAspSerGluAspThrMetSerThrGlyThrPro  985

125 GTGGTGGGCAGAACATACCTTCCCCACCCTCCACCCAAGGTGGCTCCCAT  76
       |||   ||||||  |||:::         |||:::|||
986 GlyVal...ArgThr..ProThrAsnLysPheThrAsnGly.........  997
```

```
 75 CTCTACCCCATTCCCCTGTAGAGCTCAGA.....................  47
     :::|||||||          |||||||||
998 .AsnThrPro..........GluLeuArgIleArgGluLeuAsnGlySer 1010

 47 ..................................................  47

1011 LeuAsnAlaValAsnHisLeuAlaArgGluPheAspGlnArgArgLeuAs 1027

 46 ..................................................G 46
                                                      |
1027 nPheAspGluAspAlaArgAlaIleValGluValLysLeuGlyProGlnA 1044

 45 CAAGCAGGAAACCCAAATG..AAGTGTGAAGACAGCCCAGAG          6
    ||:::       ||||||   :::   ::::::   ||||||
1044 laThr.......ProAsnGlyGlnGlnGlnGlnHisProGlu        1055
```